



ELSEVIER

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Correspondence

A fight on epistemological quicksand: Comment on the dispute between van den Besselaar et al. and Butler

1. Introduction

Two studies using different data produce different results and draw different conclusions. So what? Since we all know all indicators to be only partial indicators of what we attempt to measure, this is the expected outcome rather than a surprise. Nevertheless, I cannot possibly forego the opportunity to annoy all parties involved and part of the audience by pointing out that the debate between van den Besselaar, Heyman & Sandström (BHS) and Butler is based on the shared erroneous assumption that causal claims can be made with the measurement strategy employed.

In a nutshell, my argument is that the causal chain between performance-based funding and changed aggregate publication behavior is rather long, and that a large number of confounding variables operates at different links of the causal chain. A causal claim could be made only if all these variables were considered in conjunction, which is impossible with bibliometric methods alone.

2. Causal arguments

In a paper on evaluation-based funding in 2002, Grit Laudel, Sybille Hinze, Linda Butler and I made a statement about a possible causal relationship that used the figure later to be published by Butler and reproduced in Fig. 1 by BHS. Referring to the picture, we stated: “The timing of this productivity increase in relation to the introduction of funding formulas suggests that there is a causal relationship.” (Gläser, Laudel, Hinze, & Butler, 2002: 12). We later referred to the case studies on two Australian universities as “[providing] further support for the *assumption* that the coupling of increasing quantity and decreasing quality is due to the introduction on quantity-based funding formulas” (ibid: 14, my emphasis). Thus, we derived from Butler’s analysis the hypothesis that the introduction of the publication indicator might have caused changes in publication behavior (which was correctly quoted as a hypothesis by BHS at page 6).

Butler subsequently turned this into a causal argument. Although the texts of her publications phrase her claim as “hypothesis” for which she finds “support” (Butler, 2003b: 41, 43; 2003a: 151), titles and an abstract (Butler, 2003b: 39) contain a causal statement rather than an unproven but plausible hypothesis. The titles read “Modifying publication practices in response to funding formulas”, “Explaining Australia’s increased share of ISI publications – the effects of a funding formula based on publication counts”, and “What Happens when Funding is Linked to Publication Counts?” (Butler, 2004).

The BHS paper (van den Besselaar, Heyman, & Sandström, 2017) includes both hypothetical statements (“if output based research funding has an effect on research quality, it is positive and not negative”, ibid: abstract) and factual statements (“our aim is . . . to reanalyze the effect of the changes in the funding system in Australia. . .”, ibid: 3). They hedge their causal claim in the “conclusions and discussion” section by suggesting that “[w]e should refrain from drawing *too stark* conclusions regarding the causality of the process described. . .” It does not become clear what the ‘weak’ conclusions about causality are supposed to be. Ultimately, their claim is a causal counter-claim to Butler’s causal claim, as becomes clear in the paragraph following the ‘hedge paragraph’:

Nevertheless, the data do suggest that the new policy during the 1990s gave the Australian science system a new impulse – as funding partly became output-dependent after 1995. This initiative *did not only contribute to higher productivity*, but – as can be expected from creativity theory – *also to higher quality*. (ibid: 18, my emphasis)

But how can such a causal argument be made? With just one case, the only way to establish causality is to exclude all other possible causes by identifying the mechanism that produces changed aggregate publication and citation patterns from the introduction of the publication indicator.

3. The appropriateness of measurements

Some of the indicators used by Butler and BHS pose interesting problems for establishing causality. First, the use of Australian shares in the SCI implies an enormous complication of the causal argument because shares in international collections of publications and citations are influenced by the publication and citation activity of all other countries included in the SCI. How can one establish the extent to which a changing share of publications or citations is caused directly by a national policy under these conditions? This problem has been discussed by Hicks (2005), who argued that the US were losing their leading position in the SCI due to “the determination of Asian policy makers to strengthen their R&D systems and consequently their knowledge economies” (ibid: 8, see also Hicks, 2007: 232–238). Is it possible that increased competition for publication space in the higher-impact journals crowded out some Australian publications, which moved to lower-impact journals? In the light of changing shares, one can still argue that Australian researchers were unable or were able to maintain their relative position due to the policies introduced at home. But this is a different causal argument, namely the argument that the Australian funding formula failed to make Australian university researchers outcompete researchers in other countries who were affected by different national policies.

Second, some of the indicators used by BHS fail to support a convincing counter claim. BHS use the top 10% cited papers because “this indicator yields a direct measure of impact, while the RCI (and the Incites[®] variant of it, the NCI) is dependent of variations in low or non-cited papers, papers that do not contribute to impact” (van den Besselaar et al, 2017: 9). They thus exclude the phenomenon that is at the core of Butler’s argument – a disproportionate growth of low-impact papers. This makes it impossible to refute her claim. Acknowledging this, BHS look at the development of non-cited papers “in order to cover the various arguments by Butler” (ibid: 14, note 9). Unfortunately, this still doesn’t cover Butler’s argument because BHS only look at the poles of the spectrum of Australian publications, while Butler’s argument refers to all of them.

4. Causality and complexity

Fig. 1 illustrates the methodological problem of deriving causal claims from bibliometric analyses. Both sides assume the operation of a causal chain which produces the observed output. Butler observes that at least some universities respond to funding formulas by internally mirroring them (an observation I can confirm, see below), and suggests that these internal policies made academics produce more publications. BHS suggest that the performance-based funding of Australian universities stimulated researchers to become more productive and also increased the quality of their production. They do not further specify the causal chain.

Neither Butler nor BHS empirically investigate the assumed translation of the funding formula in university policies and of university policies into changed publication behavior. While it is not impossible to investigate these translations, it is impossible to investigate them *bibliometrically*.

Fig. 1 demonstrates that neither party is able to produce sufficient evidence for the claim that the assumed causal chain operates. My main objection goes one step further by arguing that the assumed causal chain is very unlikely to operate at all because it disregards a large number of factors that may affect publication behavior.

Both BHS and Butler are aware of the problem of confounding variables but only Butler does something about them. She discusses and excludes two alternative explanations for the changed publication pattern, namely the higher education reform at the end of the 1980s, which turned colleges of advanced education and institutes of technology into ‘new’ universities or amalgamated them with existing universities, and the rise of numbers of staff, which was considerable in the period under consideration (Butler, 2003a: 149). In order to further strengthen her claim, Butler (2003b) compares the publication output of the university sector to the outputs of the hospital and the government sector and finds that “the 1989–93 period does not mark a turning point in trends for either of these sectors” (ibid: 43).

BHS did not use any additional information of this kind. They are aware of possible confounding variables, mention a study that refuted a similar causal claim (Osuna, Cruz-Castro, & Sanz-Menendez, 2011), and list some possible confounding variables on page 18. However, they do not provide evidence for or against a causal role of any of these factors.

What other factors would need to be taken into consideration? I limit my discussion to those factors for which I can suggest plausible mechanisms (Fig. 2). For the operation of some of these mechanisms I have micro-level empirical evidence from a study on the impact of the Australian performance-based funding system conducted by Grit Laudel and myself from 2002 to 2006 (Gläser & Laudel, 2007; Gläser, Lange, Laudel, & Schimank, 2010a, Gläser, Lange, Laudel, & Schimank, 2010b).

First, it is not clear why so much steering power is ascribed to the publication indicator although it weighs only ten percent in the funding formulas. Would one not expect universities to respond most strongly to the indicator ‘external funding’, which controlled 68% of the research block grants for universities compared to the 10.3% of the publication indicator in 1996?¹ The distinct but limited role of the publication indicator is represented by a white strip in the otherwise grey box of the funding formulas.

Second, the formula-based funding occurred in the context of other higher education policies. Two of the most consequential policies were absolute reductions in government funding of higher education, which began in 1997 (Meek, 2005), and the higher education reforms discussed by Butler. I am not putting forward these policy changes as alternative explanations

¹ This can be calculated from the weights of the indicators and sums allocated provided by Butler (2003b: 40).

Download English Version:

<https://daneshyari.com/en/article/4968069>

Download Persian Version:

<https://daneshyari.com/article/4968069>

[Daneshyari.com](https://daneshyari.com)